

Running head: Physical Aggression

Teachers' Assessment of Physical Aggression with the Preschool Behavior
Questionnaire: A Multitrait-Multimethod Evaluation of Convergent and Discriminant
Validity

Jantine L. Spilt¹, Helma M.Y. Koomen¹, Reinoud D. Stoel²,
Jochem T. Thijs³, Aryan van der Leij¹

Correspondence to: J.L. Spilt, Research Institute of Child Development and
Education, University of Amsterdam, P.O. Box 94208, 1090 GE Amsterdam, the
Netherlands. Phone: +31(0)205251272; Fax: +31(0)205251200; Email:
J.L.Spilt@uva.nl

¹ Research Institute of Child Development and Education, University of Amsterdam, The Netherlands

² Netherlands Forensic Institute, The Hague, The Netherlands

³ Department of Social Science, Utrecht University, The Netherlands

Abstract

The distinctiveness of physical aggression from other antisocial behavior is widely accepted but little research has explicitly focused on young children to empirically test this assumption. A MTMM approach was employed to confirm the distinctiveness of physical aggression from nonaggressive antisocial behavior in early childhood. In addition, the convergent validity of teacher reports of physical aggression was investigated on a measure that contained age-appropriate behavior items selected from the Preschool Behavior Questionnaire (PBQ). Assessments of physical aggression versus nonaggressive antisocial behavior of 117 kindergartners with different behavior profiles were obtained using three measures varying in source and/or method: 1) teacher reports on the PBQ, 2) short interviews with teachers, and 3) classroom observations. The MTMM matrix was analyzed using confirmatory factor analysis (CFA). The results provided reasonable support for the discrimination between physical aggression and nonaggressive antisocial behavior. In addition, strong support was found for the convergent validity of teacher-reported physical aggression using PBQ items.

Keywords: physical aggression, nonaggressive antisocial behavior, convergent validity, discriminant validity, kindergarten children

Teachers' Assessment of Physical Aggression with the Preschool Behavior
Questionnaire: A Multitrait-Multimethod Evaluation of Convergent and Discriminant
Validity

It is widely accepted and empirically supported that physical aggression is a separate dimension of antisocial behavior (e.g., Cote, Vaillancourt, LeBlanc, Nagin, & Tremblay, 2006; Loeber & Hay, 1997; Nagin & Tremblay, 1999; NICHD Early Child Care Research Network, 2004). However, evidence for the assumption that physical aggression is distinctive from more generic antisocial behavior is mainly limited to middle and late childhood, whereas little research has explicitly focused on early childhood (for a review, see Moreland & Dumas, 2008). Scholars repeatedly noted that, in this age range, available measures often confound physical aggression with other troublesome behaviors (e.g., Broidy et al., 2003; Joussemet et al., 2008; Tremblay, 2000). The present study aimed to address these concerns by examining the discriminant validity of physical aggression against nonaggressive forms of antisocial behavior in a sample of kindergartners. Furthermore, we evaluated the convergent validity of teacher reports of physical aggression using an age-appropriate screening measure.

Antisocial behavior is a complex and heterogeneous concept that encompasses a variety of behaviors such as lying, indifference, bullying, fighting, theft, vandalism, and substance abuse. A major challenge is to unravel the heterogeneity of antisocial behavior and to identify different classifications. This is important as different problem behaviors, though substantially correlated, appear to have different predictive

value for physical violence. Physical aggression is more consistently related to later violence than opposition or hyperactivity (Broidy, et al., 2003; Nagin & Tremblay, 1999) and has proven a key predictor of juvenile violence (Broidy et al., 2003; Cote et al., 2006; Nagin & Tremblay, 1999; NICHD Early Child Care Research Network, 2004). Based on the current evidence in middle and late childhood, it is widely accepted that physical aggression (i.e., use of verbal threat and physical force to harm others or damage objects) is a distinct dimension of antisocial behavior that should not be confounded with other forms of antisocial behavior such as relational aggression (e.g., Crick & Grotpeter, 1995; Vaillancourt, Brendgen, Boivin, & Tremblay, 2003) and nonaggressive forms of antisocial behavior (Achenbach, Conners, Quay, Verhulst, & Howell, 1989; Frick et al., 1993; Loeber & Schmalting, 1985a, 1985b; Quay, 1987; Tackett, Krueger, Sawyer, & Graetz, 2003). These nonaggressive forms represent behaviors that are not physically or verbally abusive but do violate social norms and have the potential to be damaging to others such as lying and vandalism.

Physical aggression is associated with a wide range of concurrent and future adjustment problems, especially when it begins in early childhood (Cote et al., 2006; Joussemet et al., 2008; NICHD Early Child Care Research Network, 2004; Silver, Measelle, Armstrong, & Essex, 2005; Tremblay, 2004). Severe problems are observed in both the social and cognitive domain, such as social rejection and academic underachievement (e.g., Dodge et al., 2003; Stipek & Miles, 2008). Moreover, about 4.4 to 16.6% of the young children persist in the use of physical aggression up to adolescence and adulthood (for a review, see Nagin & Tremblay, 2005). Early intervention is therefore crucial to counteract developmental trajectories of chronic aggression (e.g., Webster-Stratton, 1997). However, scholars have repeatedly noted

that, for young children, physical aggression is often confounded with other forms of antisocial behavior, which compromises predictive validity and the detection of children in need of intervention (Broidy et al., 2003; Joussemet et al., 2008; Tremblay, 2000). In this age range rigorous tests of the assumption that physical aggression is a distinct dimension of antisocial behavior are sparse. It is possible that early antisocial behavior is a unidimensional construct that may become more specific and differentiated with age (cf. Weis, Lovejoy, & Lundahl, 2005). When this would be confirmed, the call for research and screening practices that focus on pure physical rather than generic antisocial behavior in young children may be pointless and perhaps even counterproductive (e.g., decreasing predictive validity and screening accuracy).

Recent research supports the distinctiveness of physical aggression from nonaggressive antisocial behavior in kindergarten children (Spilt, Koomen, Thijs, Stoel, & Van der Leij, 2010). Teacher reports of antisocial behavior were obtained using the Preschool Behavior Questionnaire (PBQ; Behar, 1977). This screening measure is specifically developed for young children and contains a heterogeneous pool of age-appropriate items of antisocial behavior. The checklist is easy to complete and widely used in international research including school-based studies (e.g., Drugli & Larsson, 2006; Goossens, Bokhorst, Bruinsma, & Van Boxtel, 2002; Silver et al., 2005; Tremblay, Desmarais-Gervais, Gagnon, & Charlebois, 1987). The results of Spilt and colleagues (2010) showed that physical aggression could be distinguished from other troublesome behaviors such as not sharing, blaming others, and showing sneaky behavior. The distinction was tested and cross-validated for both boys and girls using CFA in two large samples. In addition, differential associations with

internalizing behavior and different outcomes with respect to gender differences in both mean levels and measurement issues (i.e., factorial invariance) were reported. It was suggested that a subset of PBQ items could be used as a screening measure to obtain teacher reports of physical aggression (Spilt et al., 2010). Teachers are widely acknowledged as valuable sources of information on children's social-behavioral functioning. Especially their reports on dimensions of externalizing behavior show strong validity (Konold & Pianta, 2007). Given the vital role of teachers in the detection of at-risk children, it is crucial to evaluate the convergent validity of teacher measures of physical aggression that are specifically designed for young children.

A widely-used approach to evaluate convergent validity is the multi-trait multi-method model (MTMM; Campbell & Fiske, 1959). In this approach, trait variance is isolated from method or source variance and from unique variance (i.e., random variance). Trait effects represent systematic variance in a manifest variable associated with a particular latent trait. Method effects refer to systematic variance that is specific to a certain method or source. Significant trait variance indicates convergent validity, whereas substantial method variance weakens the support for convergent validity. Furthermore, the MTMM model allows for an estimation of the true association between constructs. The degree of trait covariance is an indicator of discriminant validity (i.e., the degree to which constructs are distinct). Thus, in addition to previous CFA research (Spilt et al., 2010), the MTMM approach offers an alternative way to test and confirm the distinctiveness of physical aggression from other nonaggressive-antisocial behaviors.

The first goal of the current investigation was to provide further evidence for the separate assessment of physical aggression in young children by demonstrating

discriminant validity in a behaviorally-diverse sample of kindergartners. The second goal was to evaluate the convergent validity of teacher-reported physical aggression using PBQ items. Teacher ratings of physical aggression were evaluated against two alternative measures varying in source and/or method, namely teachers' free descriptions of children's social behavior and independent classroom observations. Based on previous research, we anticipated moderately large convergence between PBQ ratings and free descriptions of physical aggression (cf. Thijs, Koomen, De Jong, Van der Leij, & Van Leeuwen, 2004). Modest convergence with observations was predicted (cf. Hoge, Meginbir, Khan, & Weatherall, 1985; Ostrov & Keating, 2004; Rubin, Moller, & Emptage, 1987). Furthermore, though we expected a substantially large amount of shared-trait variance between physical aggression and nonaggressive antisocial behavior, because both represent dimensions of antisocial behavior, some degree of differentiation was anticipated based on previous research (cf. Konold & Pianta, 2007; Spilt et al., 2010; Tackett et al., 2003).

Method

Participants

The sample was part of a larger investigation into teachers' cognitions and practices toward children with different behavior profiles (Thijs, 2005). Data were available for 117 children from 39 kindergarten classrooms of regular elementary schools located in different rural and nonrural parts of the Netherlands. Written informed consent was obtained from the parents. Teachers ($N = 39$, two men) had reported on 858 children (414 girls) with a mean age of 67.5 months ($SD = 8.4$). Three children with different behavior profiles were selected from each classroom,

without teachers being informed about the selection criteria. The selection of the children was based on their PBQ scores in comparison to the scores of their classmates. Children with the highest ratings on Social Withdrawal but not above the mean on Externalizing Behavior were selected for the withdrawn subgroup. In contrast, children with the highest ratings on Externalizing Behavior but not above the mean on Social Withdrawal and Internalizing Behavior were assigned to the externalizing subgroup. The average group consisted of children with scores close to but not above the class mean on these three scales. The withdrawn, externalizing, and average subgroup included 59.0%, 33.3%, and 69.2% girls, respectively. Considering the generally low prevalence of behavior problems in regular classrooms, with this selection procedure we guaranteed that children with a variety of behaviors were included to obtain an adequate sample for validation purposes.

Measures

Questionnaire. The Preschool Behavior Questionnaire (PBQ; Behar, 1977) is a checklist that contains age-appropriate descriptions of behavior problems of young children. Items are rated on a 4-point Likert scale, ranging from 1 (absolutely not characteristic) to 4 (very characteristic). Teachers completed a Dutch version that has shown strong evidence of reliability and validity (Goossens, Dekker, Bruinsma, & De Ruyter, 2000). The Externalizing scale has shown high internal consistency ($\alpha \geq .91$), test-retest stability ($rs \geq .84$), and interrater agreement ($\alpha = .91$) in a community and clinical sample. The validity has been supported by concurrent and predictive associations with parallel teacher-rated measures of adjustment, peer evaluations of aggression, and children's sociometric status (Goossens et al., 2002; Goossens et al.,

2000). In addition, the scale has been shown to discriminate between a clinical and community sample (Goossens et al., 2000).

The Externalizing scale contains 14 items, including 8 antisocial behavior items. The distinction between Physical Aggression (PA: 4 items: e.g., 'Kicks or hits' and 'Bullies'; $\alpha=.90$) and Nonaggressive Antisocial Behavior (NAB: 4 items: e.g., 'Sneaky' and 'Blames others'; $\alpha=.88$) has been supported by confirmatory factor analytic research in two independent samples of kindergarten boys and girls (Spilt et al., 2010).

To select the target children, we used the broad-band scales for Externalizing and Internalizing behavior (15 items, $\alpha=.89$), and the aggregated measure for Social Withdrawal (8 items, $\alpha=.84$) containing the items of the subscales Social Inhibition (e.g., 'Shy or timid') and Solitary Behavior (e.g., 'Somewhat on his/her own'). Discriminant and convergent validity of those subscales has been supported (Thijs et al., 2004).

Classroom observations. Videotapes were made during a small-group activity in the natural ecology of the classroom that lasted on average 15 minutes ($M=14.2$; $SD=2.3$). The three selected children worked together on a large jigsaw puzzle. Teachers were given no specific instructions but were merely asked to attend the activity as they normally would do. Videotapes of the puzzle situation were coded for the occurrence of PA and NAB. Occurrences of antisocial behavior were scored only for those periods in which the teacher was absent.

Observations of PA involved the frequency of verbal provocation, mean gestures, physical aggression, and object aggression. Observed NAB included violations of social norms (i.e., not sharing, refusal to cooperate, lying, ignorance or exclusion of

other children, mean gestures or comments behind ones back, blaming others). Total frequency counts were divided by the total number of observation minutes.

Two independent observers, who were unaware of children's behavior profile and study purposes, scored the videotapes. Intraclass correlations (ICC) for absolute agreement based on average measures were .90 and .89 for PA and NAB, respectively. Mean scores of the two observers were used in the analyses.

Interviews. Teachers were interviewed about each of the selected children to obtain free descriptions of teachers about children's social behavior. In each interview, children were discussed in random order. An open-ended question was asked: "How would you describe the child's general social behavior, as compared to his/her age mates?" The interviewer did not refer to PA and NAB. Interviews were conducted by telephone and recorded on audio tape. They took place on average 5.3 weeks after the PBQ was completed ($SD=3.0$ weeks).

Four independent raters coded the interviews. Raters were blind to selection criteria and study purposes. The rating form included 7 items referring to problem behaviors of which two items, representing PA and NAB, were used in the current study. Eight items addressed the presence of positive behaviors to reduce the chance that raters would be exclusively focused on problem behavior. Items were rated on a 4-point Likert scale, ranging from 1 (*not characteristic at all*) to 4 (*very characteristic*). ICCs for consistency in agreement based on average measures were .80 and .61 for PA and NAB, respectively. Mean scores of the four raters were utilized in the analyses.

Statistical analyses

The MTMM model was estimated by means of confirmatory factor analysis (CFA) using the Mplus program (Muthén & Muthén, 1998-2005). Analyses were conducted on the observed data under the assumption of Missing At Random. To account for the non-independence and the non-normal distribution of the data, we used the cluster option of Mplus with TYPE=Complex routine (children nested in teachers) and the robust maximum likelihood method MLR. The MLR chi-square test statistic was used to evaluate overall model fit, also referred to as the Yuan-Bentler chi-square statistic. Additional fit indexes were examined because the chi-square is highly sensitive to sample size. The model fit is considered acceptable when $CFI \geq .95$, $SRMR \leq .08$, and $RMSEA \leq .06$ (Hu & Bentler, 1999). Because the Yuan-Bentler chi-square cannot be used for difference testing of nested models, the Satorra-Bentler scaled chi-square difference test was used (Satorra & Bentler, 2001).

Results

Preliminary analyses

Ten cases with extreme scores (Z-scores above 2) were detected in the observations of PA and NAB. Extreme scores were mostly given to children within the same small group. In one group, the teacher did not attend the children for a relatively long period of 10 minutes. Physical fighting escalated during this time resulting in extreme ratings for all three children, whereas in other groups no such substantial fights were observed. In another group, four instead of three children were working together. Furthermore, one child received a high rating on physical aggression because the child was repeatedly throwing with things. Notably, extreme scores were observed across all children independent of children's behavioral profiles.

These 10 cases were considered univariate outliers and excluded from subsequent analyses.

Descriptive statistics

Table 1 presents the descriptive statistics. Substantial variance was observed for all variables. The correlation between PA and NAB varied across measurement methods. A high correlation was found among questionnaire reports, whereas interview and observational ratings yielded a mediocre and modest correlation, respectively. Questionnaire reports of PA were related to interview ratings of both PA and NAB but seemed somewhat stronger associated with interview PA than NAB. Furthermore, questionnaire reports of PA were positively correlated with observed PA but not with observed NAB. All correlations were in the expected directions.

Multitrait-Multimethod model

Several nested models were evaluated, starting with a model without method factors. Method factors were specified to be uncorrelated with the other factors in the model. To retain degrees of freedom, the unstandardized loadings on the method factors were constrained to be equal for variables measured by the same method. Note that the error variances were not constrained resulting in unequal standardized loadings (see Figure 1).

The final most parsimonious model presented in Figure 1 contained two trait and two method factors. The two method factors represented unique source variance (i.e., teacher and independent observer). The fit of this model was good: χ^2 (6, $N=107$) = 3.28, $p=ns$; RMSEA=.00; SRMR=.03; CFI=1.00, and significantly better than a

model without method factors: $\Delta\chi^2=12.88$, $\Delta df=2$, $p<.001$. In addition, a two-trait model with three method factors (independent observation, teacher-questionnaire, and teacher-interview with the latter two allowed to be correlated) did not provide a better fit to the data than the final model: $\Delta\chi^2=.20$, $\Delta df=2$, $p=ns$.

Discriminant validity. The covariance between PA and NAB was reasonably high with 65.6% shared-trait variance. To provide a more stringent test of the distinctiveness of the traits, the final model was compared to a similar model with one trait factor. This latter model fitted the data significantly worse ($\Delta\chi^2=9.40$, $\Delta df=1$, $p<.01$).

Convergent validity. Marginally to strongly significant trait variance was observed in the measures of PA, providing support for convergence among different methods and sources. Teacher-questionnaire reports showed large trait variance, whereas teacher interviews showed moderately large trait variance. Observer ratings demonstrated modest trait variance that was marginally significant.

Method effects. All measures demonstrated the presence of method effects. In support of convergent validity, method effects in teacher-questionnaire reports of PA and NAB were substantially smaller than trait effects. In teacher interviews and observer ratings of PA and NAB, equally large or larger method effects were observed in comparison to trait effects.

Discussion

The distinctiveness of physical aggression from other antisocial behavior is widely accepted but little research has explicitly focused on young children to empirically test this assumption. In addition, physical aggression in early childhood is often

confounded with other forms of antisocial behaviors, which limits understanding of the developmental dynamics of physical aggression and compromise screening as scholars have repeatedly noted (e.g., Broidy et al., 2003; Joussemet et al., 2008; Tremblay, 2000). This study addressed these concerns in two important ways. First, using a MTMM approach, the discriminant validity of physical aggression against nonaggressive antisocial behavior was largely confirmed in a behaviorally-diverse sample of kindergartners. Second, the convergent validity of teacher reports of physical aggression on the PBQ was supported.

In the current study, assessments of physical aggression and nonaggressive antisocial behavior were obtained using three measures that differed in method and/or source. In the MTMM approach, method and/or source effects are isolated such that true relations among traits can be estimated. A considerable between-trait association between physical aggression and nonaggressive antisocial behavior was found. This is not surprising because the constructs represent dimensions of the same higher-order construct of antisocial behavior. Differences in antisocial strategies may be subtle and not easily observable for teachers. Also, problems tend to co-occur and many aggressive children are likely to demonstrate other forms of antisocial behavior at times as well. For comparison, a between-trait correlation of .88 was found between the subscales Delinquency and Aggression of the Child Behavior Checklist in a multitrait-multisource design including parents and teachers (Konold & Pianta, 2007). But importantly, the overlap between physical aggression and nonaggressive antisocial behavior was not complete. The estimation of 34.4% unique trait variance indicates some differentiation as well. Moreover, the comparison of the one-trait with the two-trait model offers a more stringent test and also supported the distinctiveness

of the constructs. This finding converges with previous research using confirmatory factor analysis in two large samples of kindergartners (Spilt et al., 2010). The divergence between the two constructs may be embedded in differences in etiology and risk factors such as parental traits and psychopathology, parenting skills and monitoring, stability, and heritability (Eley, Lichtenstein, & Stevenson, 1999; Monuteaux, Fitzmaurice, Blacker, Buka, & Biederman, 2004; Nigg & Hinshaw, 1998; Patrick, Snyder, Schrepferman, & Snyder, 2005; Stanger, Achenbach, & Verhulst, 1997). In addition, physical aggression could lead to different developmental outcomes. Prospective research indicated that physically aggressive children especially are at at-risk of juvenile violence (e.g., Nagin & Tremblay, 1999; NICHD Early Child Care Research Network, 2004), whereas nonaggressive antisocial behavior appeared more predictive of nonviolent delinquency (Nagin & Tremblay, 1999). Together with the current results, these findings emphasize that the distinctiveness of physical aggression from more generic antisocial behavior, though moderate, should not be ignored.

Support was found for the convergent validity of teacher-reported physical aggression on the PBQ. PBQ reports of physical aggression contained a relatively small amount of source variance in comparison to trait variance. As predicted, moderately large convergence was found with teachers' free descriptions of children's social behavior (cf. Thijs et al., 2004). Also consistent with our expectations and previous research, modest teacher-observer convergence was found (e.g., McNeilly-Choque, Hart, Robinson, Nelson, & Olsen, 1996). The discrepancy between observer and teacher ratings may not simply represent 'noise' but is also likely to reflect true differences that result from the situation in which the judgments are made. Teachers'

judgments were rooted in many daily observations and interactions with the children across various situations and school contexts. In contrast, observer ratings were restricted to occurrences of children's antisocial behavior in a specific small-group peer situation when the teacher was not present as well as to a limited time frame (Ladd & Profilet, 1996; Lakes & Hoyt, 2008). Notwithstanding the divergence, the results do suggest that teacher ratings of physical aggression are to some extent related to actual physical aggression in a specific context, as has been found in other observational studies (McEvoy, Estrem, Rodriguez, & Olson, 2003; McNeilly-Choque et al., 1996; Ostrov & Keating, 2004).

Several qualifications of the study should be considered. Data were obtained in a behaviorally-diverse subsample of kindergarten children. Analyses in more homogeneous or randomly-selected samples could have yielded different results. Although the observations were conducted in a natural activity in the classroom, supporting ecological validity, the observations were restricted to one specific small-group activity and generalizability to other situations may be low (Lakes & Hoyt, 2008). The specific peer context, the task used, and teacher characteristics may have affected children's use of aggression. Multiple observations capturing behavior across school activities (e.g., free play) are recommended in future research.

For school psychologists, teachers are an imperative source of information on children's social-behavioral functioning. It is therefore of crucial importance that teacher measures routinely undergo critical evaluation. The current study addressed the repeated call of scholars for a screening strategy focused on physical aggression rather than more generic antisocial behavior to increase predictive accuracy. The results support the use of the PBQ as a brief screening measure to obtain information

from teachers about young children's physically aggressive behavior. Note that the scale is not developed for diagnostic use. Practitioners are urged to employ multiple sources, and assess physical aggression in the context of other key variables such as family adversity (e.g., Patrick et al., 2005).

In conclusion, the present study evaluated the discriminant and convergent validity of teacher reports of physical aggression using an age-appropriate measure for young children derived from the Preschool Behavior Questionnaire (PBQ). Support was found for the separate assessment of physical aggression from nonaggressive antisocial behavior. The finding of substantially more trait than source and method variance in PBQ reports supported convergent validity.

References

- Achenbach, T. M., Conners, C. K., Quay, H. C., Verhulst, F. C., & Howell, C. T. (1989). Replication of empirically derived syndromes as a basis for taxonomy of child/adolescent psychopathology. *Journal of Abnormal Child Psychology*, 17, 299-323.
- Behar, L. B. (1977). The Preschool Behavior Questionnaire. *Journal of Abnormal Child Psychology*, 5, 265-275.
- Bentler, P. M., & Chou, C.-P. (1987). Practical Issues in Structural Modeling. *Sociological Methods Research*, 16(1), 78-117.
- Broidy, L. M., Nagin, D. S., Tremblay, R. E., Bates, J. E., Brame, B., Dodge, K. A., et al. (2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psychology*, 39, 222-245.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
- Cote, S., Vaillancourt, T., LeBlanc, J., Nagin, D., & Tremblay, R. (2006). The development of physical aggression from toddlerhood to pre-adolescence: A nation wide longitudinal study of canadian children. *Journal of Abnormal Child Psychology*, 34, 68-82.
- Crick, N. R., & Grotpeter, J. K. (1995). Relational aggression, gender, and social-psychological adjustment. *Child Development*, 66(3), 710-722.
- Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., et al. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development*, 74, 374-393.
- Drugli, M., & Larsson, B. (2006). Children aged 4-8 years treated with parent training and child therapy because of conduct problems: generalisation effects to day-care and school settings. *European Child & Adolescent Psychiatry*, 15, 392-399.
- Eley, T. C., Lichtenstein, P., & Stevenson, J. (1999). Sex differences in the etiology of aggressive and nonaggressive antisocial behavior: Results from two twin studies. *Child Development*, 70, 155-168.

- Frick, P. J., Lahey, B. B., Loeber, R., Tannenbaum, L., Van Horn, Y., Christ, M. A. G., et al. (1993). Oppositional defiant disorder and conduct disorder: A meta-analytic review of factor analyses and cross-validation in a clinic sample. *Clinical Psychology Review, 13*, 319-340.
- Goossens, F. A., Bokhorst, K., Bruinsma, C., & Van Boxtel, H. W. (2002). Judgments of aggressive, withdrawn and prosocial behavior: Perceived control, anger, pity and sympathy in young Dutch children. *Journal of School Psychology, 40*, 309-327.
- Goossens, F. A., Dekker, P., Bruinsma, C., & De Ruyter, P. A. (2000). *De Gedrags Vragenlijst voor Peuters en Kleuters: Factor structuur, betrouwbaarheid en validiteit. [Behavior questionnaire for 2- to 6-year-olds: factor structure, reliability, and validity]*. Amsterdam: Vrije Universiteit.
- Hoge, R. D., Meghinbir, L., Khan, Y., & Weatherall, D. (1985). A multitrait-multimethod analysis of the preschool behavior questionnaire. *Journal of Abnormal Child Psychology, 13*, 119-127.
- Joussemet, M., Frank Vitaro, F., Barker, E. D., Côté, S., Nagin, D. S., Zoccolillo, M., et al. (2008). Controlling parenting and physical aggression during elementary school. *Child Development, 79*, 411-425.
- Konold, T. R., & Pianta, R. C. (2007). The influence of informants on ratings of children's behavioral functioning: A latent variable approach. *Journal of Psychoeducational Assessment, 25*, 222-236.
- Ladd, G. W., & Profilet, S. M. (1996). The Child Behavior Scale: A teacher-report measure of young children's aggressive, withdrawn, and prosocial behaviors. *Developmental Psychology, 32*, 1008-1024.
- Lakes, K. D., & Hoyt, W. T. (2008). What sources contribute to variance in observer ratings? Using generalizability theory to assess construct validity of psychological measures. *Infant and child development, 17*, 269-284.
- Loeber, R., & Hay, D. (1997). Key issues in the development of aggression and violence from childhood to early adulthood. *Annual Review of Psychology, 48*, 371-410.

- Loeber, R., & Schmalting, K. B. (1985a). Empirical evidence for overt and covert patterns of antisocial conduct problems: A metaanalysis. *Journal of Abnormal Child Psychology*, 13, 337-353.
- Loeber, R., & Schmalting, K. B. (1985b). The utility of differentiating between mixed and pure forms of antisocial child behavior. *Journal of Abnormal Child Psychology*, 13, 315-335.
- McEvoy, M. A., Estrem, T. L., Rodriguez, M. C., & Olson, M. L. (2003). Assessing relational and physical aggression among preschool children: Intermethod agreement. *Topics in Early Childhood Special Education*, 23, 53-63.
- McNeilly-Choque, M. K., Hart, C. H., Robinson, C. C., Nelson, L. J., & Olsen, S. F. (1996). Overt and relational aggression on the playground: correspondence among different informants. *Journal of research in childhood education*, 11, 47-67.
- Monuteaux, M. C., Fitzmaurice, G., Blacker, D., Buka, S. L., & Biederman, J. (2004). Specificity in the familial aggregation of overt and covert conduct disorder symptoms in a referred attention-deficit hyperactivity disorder sample. *Psychological Medicine*, 34, 1113-1127.
- Moreland, A. D., & Dumas, J. E. (2008). Categorical and dimensional approaches to the measurement of disruptive behavior in the preschool years: A meta-analysis. *Clinical Psychology Review*, 28, 1059-1070.
- Muthén, L. K., & Muthén, B. O. (1998-2005). Mplus (Version 3.13). Los Angeles: Muthén & Muthén.
- Nagin, D. S., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development*, 70, 1181-1196.
- Nagin, D. S., & Tremblay, R. E. (2005). What has been learned from group-based trajectory modeling? Examples from physical aggression and other problem behaviors. *The ANNALS of the American Academy of Political and Social Science*, 602, 82-117.
- NICHD Early Child Care Research Network. (2004). Trajectories of physical aggression from toddlerhood to middle childhood. *Monographs of the Society for Research in Child Development*, 96(4), 1-128.

- Nigg, J. T., & Hinshaw, S. P. (1998). Parent personality traits and psychopathology associated with antisocial behaviors in childhood Attention-Deficit Hyperactivity Disorder. *Journal of Child Psychology and Psychiatry*, 39, 145-159.
- Ostrov, J. M., & Keating, C. F. (2004). Gender differences in preschool aggression during free play and structured interactions: An observational study. *Social Development*, 13, 255-277.
- Patrick, M. R., Snyder, J., Schrepferman, L. M., & Snyder, J. (2005). The joint contribution of early parental warmth, communication and tracking, and early child conduct problems on monitoring in late childhood. *Child Development*, 76, 999-1014.
- Quay, H. C. (1987). Patterns of delinquent behavior. In H. C. Quay (Ed.), *Handbook of juvenile delinquency* (pp. 118-138). New York: John Wiley & Sons.
- Rubin, K. H., Moller, L., & Emptage, A. (1987). The Preschool Behaviour Questionnaire: A useful index of behaviour problems in elementary school-age children? *Canadian Journal of Behavioural Science*, 19, 86-100.
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher-child relationship during the school transition. *Journal of School Psychology*, 43, 39-60.
- Spilt, J. L., Koomen, H. M. Y., Thijs, J., Stoel, R. D., & Van der Leij, A. (2010). Teachers' assessment of antisocial behavior in kindergarten: Physical aggression and measurement bias across gender. *Journal of Psychoeducational Assessment*, 28, 129-138.
- Stanger, C., Achenbach, T. M., & Verhulst, F. C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Development and Psychopathology*, 9, 43-58.
- Stipek, D., & Miles, S. (2008). Effects of aggression on achievement: Does conflict with the teacher make it worse? *Child Development*, 79, 1721-1735.

- Tackett, J. L., Krueger, R. F., Sawyer, M. G., & Graetz, B. W. (2003). Subfactors of DSM-IV conduct disorder: Evidence and connections with syndromes from the Child Behavior Checklist. *Journal of Abnormal Child Psychology*, 31, 647-654.
- Thijs, J. T. (2005). *Teachers' relationships and interactions with socially inhibited children* University of Amsterdam, SCO-Kohnstamm Instituut
- Thijs, J. T., Koomen, H. M. Y., De Jong, P. F., Van der Leij, A., & Van Leeuwen, M. G. P. (2004). Internalizing behaviors among kindergarten children: measuring dimensions of social withdrawal with a checklist. *Journal of Clinical Child and Adolescent Psychology*, 33, 802-812.
- Tremblay, R. E. (2000). The development of aggressive behaviour during childhood: What have we learned in the past century? *International Journal of Behavioral Development*, 24, 129-141.
- Tremblay, R. E. (2004). Decade of behavior distinguished lecture: Development of physical aggression during infancy. *Infant Mental Health Journal*, 25, 399-407.
- Tremblay, R. E., Desmarais-Gervais, L., Gagnon, C., & Charlebois, P. (1987). The Preschool Behaviour Questionnaire: Stability of its factor structure between cultures, sexes, ages and socioeconomic classes. *International Journal of Behavioral Development*, 10, 467-484.
- Vaillancourt, T., Brendgen, M., Boivin, M., & Tremblay, R. E. (2003). A longitudinal confirmatory factor analysis of indirect and physical aggression: Evidence of two factors over time? *Child Development*, 74, 1628-1638.
- Webster-Stratton, C. (1997). Early intervention for families of preschool children with conduct problems. In M. J. Guralnick (Ed.), *The effectiveness of early intervention* (pp. 429-453). Baltimore: Brookes.
- Weis, R., Lovejoy, M. C., & Lundahl, B. W. (2005). Factor structure and discriminative validity of the Eyberg Child Behavior Inventory with young children. *Journal of Psychopathology and Behavioral Assessment*, 27, 269-278.

Tables

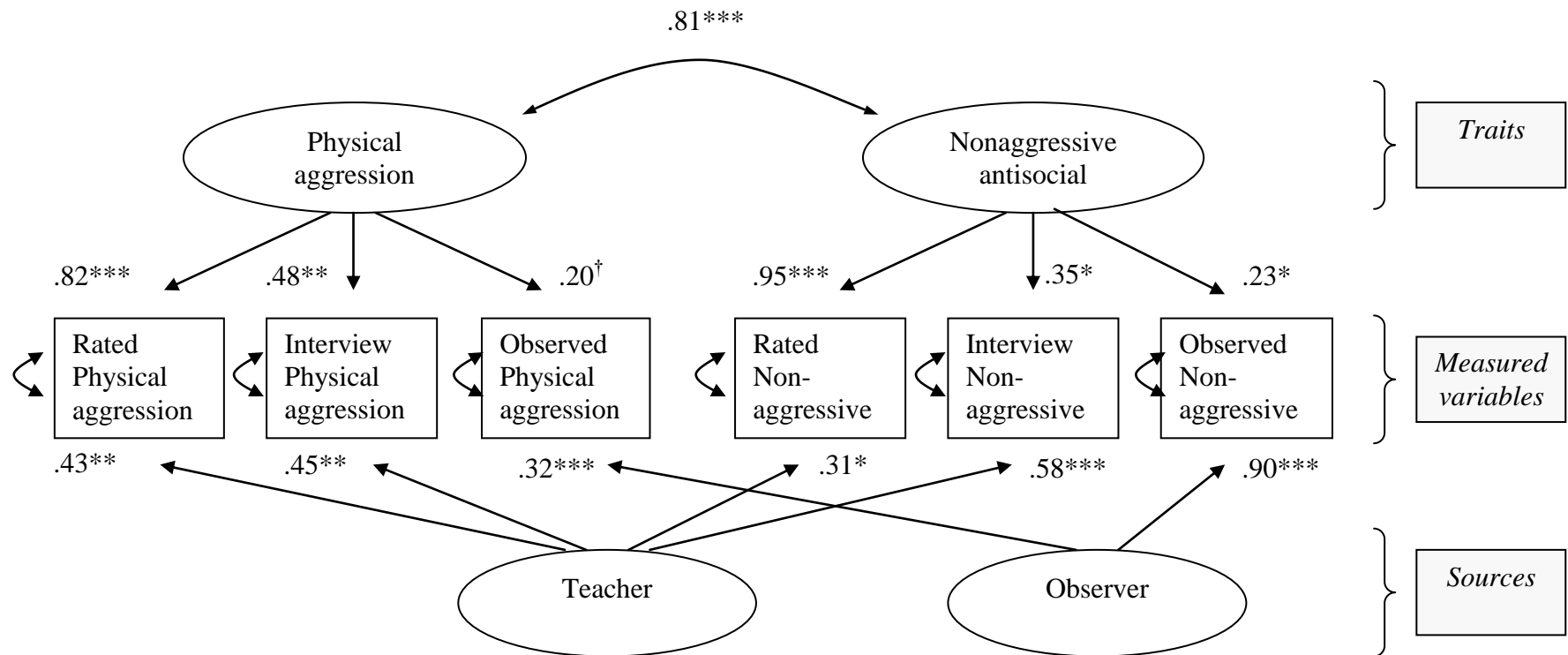
Table 1.

Zero-order correlation matrix with means and standard deviations (N=107)

Measured variables	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
<i>Questionnaire</i>								
1. Physical aggression	1.30	0.53	-	.7665**	.5746**	.4376**	.1905*	.1364
2. Nonaggressive antisocial	1.48	0.72		-	.5184**	.5199**	.1208	.1973*
<i>Interview</i>								
3. Physical aggression	1.30	0.50			-	.4503**	.2078*	.0563
4. Nonaggressive antisocial	1.21	0.39				-	.0867	.0878
<i>Observation</i>								
5. Physical aggression ^a	0.37	0.29					-	.3071**
6. Nonaggressive antisocial ^a	0.07	0.10						-

Note 1. * $p < .05$, ** $p < .01$ (one-tailed)*Note 2.* ^a= rate per minute

Figure 1.
Multitrait-multimethod model with standardized estimates



Note: $*p < .05$, $**p < .01$, $***p < .001$, $^{\dagger}p = .065$ (one-tailed)

This is an Author's Accepted Manuscript of an article published in *Journal of Psychoeducational Assessment*, 2011, 29, 407-417. © SAGE Publications, the final version is available online at <http://dx.doi.org/10.1177/0734282910388881>

